Deepfake Detection: A Systematic Literature Review

Year: 2022 | Citations: 294 | Authors: M. Rana, M. N. Nobi, B. Murali, A. Sung

---

Abstract

Over the last few decades, rapid progress in AI, machine learning, and deep learning has resulted in new techniques and various tools for manipulating multimedia. Though the technology has been mostly used in legitimate applications such as for entertainment and education, etc., malicious users have also exploited them for unlawful or nefarious purposes. For example, high-quality and realistic fake videos, images, or audios have been created to spread misinformation and propaganda, foment political discord and hate, or even harass and blackmail people. The manipulated, high-quality and realistic videos have become known recently as Deepfake. Various approaches have since been described in the literature to deal with the problems raised by Deepfake. To provide an updated overview of the research works in Deepfake detection, we conduct a systematic literature review (SLR) in this paper, summarizing 112 relevant articles from 2018 to 2020 that presented a variety of methodologies. We analyze them by grouping them into four different categories: deep learning-based techniques, classical machine learning-based methods, statistical techniques, and blockchain-based techniques. We also evaluate the performance of the detection capability of the various methods with respect to different datasets and conclude that the deep learning-based methods outperform other methods in Deepfake detection.